

# Bayesian Framework for Mobility Pattern Discovery Using Mobile Network Events

Somayeh Danafar  
Institute of Applied Simulation, ZHAW  
Wädenswil, Switzerland

Michal Piorkowski  
Swisscom  
Bern, Switzerland

Krzysztof Kryszczuk  
Institute of Applied Simulation, ZHAW  
Wädenswil, Switzerland

**Abstract**—Understanding human mobility patterns is of great importance for planning urban and extra-urban spaces and communication infrastructures. The omnipresence of mobile telephony in today's society opens new avenues of discovering the patterns of human mobility by means of analyzing cellular network data. Of particular interest is analyzing passively collected Network Events (NEs) due to their scalability. However, mobility pattern analysis based on network events is challenging because of the coarse granularity of NEs. In this paper, we propose network event-based Bayesian approaches for mobility pattern recognition and reconstruction, mode of transport recognition and modeling the frequent trajectories.

## I. INTRODUCTION

Every day, billions of people use mobile devices, producing large volumes of monitoring signaling events resulting from the usage of the mobile terminals, which are recorded by the mobile operators. The primary reason for collecting these events, referred to “Network Events” (NEs), is quality assurance of communication services.

Wide-spread popularity of cellular phones and recorded NEs of such large group of people turn NEs as a useful source of information to characterize human behavior and mobility over long time periods [1]. The analyses of users' activities enable a network operator to provide a wide variety of services, including location-based services [2], [3], traffic management and prediction [4], and public transport planning [5]. Mobile service providers can collect NEs passively, providing smart network-based services at negligible extra costs. The automatic, passive NE collection requires no extra battery-draining applications, or dedicated hardware components installed on the mobile phone. No explicit user cooperation is required either. This reasoning motivates network operators in further analysis of NEs in the context of user mobility insights.

The network supports a number of elaborate telecommunication protocols (<http://www.3gpp.org/specifications>). This implies a high complexity for the signaling traffic, *e.g.* telco's infrastructure consists of a radio access network (RAN) and a core network (CN). The radio communication takes place between a mobile device and a base station within RAN, serving one or more radio cells, which then carries the traffic (voice & data) via fixed networks to/from CN. Radio cells (Base Transceiver Stations (BTS)) are the smallest spatial entities in the cellular network. Depending on the radio bearer, they can be classified as 2G (GSM/EDGE), 3G (UMTS/HSPA) or 4G (LTE). A passive monitoring system collects signaling

events (NEs) either from the links between the RAN and CN parts of the network, covering all 2G, 3G and 4G (A, Gb, IuPS, IuCs and S1-MME interfaces) or from RAN network elements directly (GPEH and LTE CTR).

The location of a mobile phone is approximated using the position of the BTS that user was communicating with at the time, creating an opportunity of designing a GPS-free, network-based, mobility mapping system. However, localizing a mobile terminal solely using the NEs is a challenging task, for following reasons: (1) The location of the active BTS can be far away (often kilometers) from the real location of the user. (2) The spatial resolution of the NEs is dependent on the coverage area of BTSs, and how densely BTSs are distributed. Therefore, spatial resolution of NEs is much lower than GPS data. (3) Finally, the temporal resolution of NEs is irregular and highly dependent on network activity and the frequency of phone activities. The temporal gap between two consecutive phone activities may vary between few seconds up to hours.

Considering aforementioned problems, we propose Bayesian methods for trajectory reconstruction, frequent trajectory modeling and the mode of transport recognition.

This paper is organized as follow. section II compares related literature in the context of mobility mapping analysis and elaborates our contributions respectively. We briefly explain the used database in section III. Proposed methods of trajectory reconstruction is presented in section IV, and for modeling the most representing NEs associated to frequently traversed trajectories in section V. Mode of transport recognition is discussed in section VI.

## II. RELATED WORK AND OUR CONTRIBUTIONS

A fundamental part of research in mobility data analysis is extracting spatial and/or temporal patterns which describe individual/group movements, and defining mobility models for frequent behaviors [6]. Giannotti *et al.* in [7] defined mobility patterns as groups of trajectories representing common behaviors such as visiting same regions in the same time interval and in the same order. They modeled the mobility patterns statistics with sets of distributions within the context of specific applications. Most literature in network-based mobility mapping, such as [7]–[9] focused on macro level mobility pattern analysis (in which movement of huge amount of people together is considered). The novelty of the present study is a focus on micro-, NE-based mobility pattern analysis,

which enables phone operators to offer tailored services to individual and public customers.

**1) Trajectory Reconstruction.** Fillekes [10] developed individual trajectory reconstruction techniques based on Call Detail Record (CDR) data. She investigated different map matching approaches, where CDR fixes are assigned to nodes of a routable geographical map and the user's path are correlated on this map. The route connecting consecutive CDR nodes are extracted by a sequential shortest-path analysis. Thereafter spatial and spatio-temporal similarity measures are used to compare the reconstructed trajectories with relevant GPS trajectories as ground truth paths.

**Contribution.** CDR data only include voice call and SMS activities. Capturing spatio-temporal characteristics of individual mobility using CDRs is difficult if users do not constantly use their phone [11]. Rather than working with CDRs, we use the network monitoring data (NEs). Besides voice call and SMS events, NEs contain also low level telecommunication protocol events, such as radio channel switching, handovers, and data connection setup. Therefore, NEs have substantially better spatio-temporal resolution than CDRs.

Comparing to [10] that uses a time-consuming sequential shortest-path analysis for all possible paths between designated start and destination, we extract shortest and fastest path analysis based on different means of transport between the first and last NEs. Thereafter our Bayesian model infer which extracted route is the most probable route traversed by user given observed set of NEs. We reconstruct the trajectory based on the selected route and statistically analyze how far these trajectories are located comparing to the ground truth (GPS) trajectories.

**2) Trajectory Modeling.** Smoreda *et al.* [12] used NEs information of individual users to extract and identify their common daily short distance mobility motifs. Even though millions of motifs may exist, they have shown seven simple motifs describe eighty percent of population mobility patterns. In [13] authors have classified mobility motifs in two categories: a home-workplace motif, and others. They first identify the anchor locations (*e.g.* home and workplace) for each user, thereafter extract the shortest geographical route between these specified origin and destination. If the trajectory patterns of a user matches with the extracted route, corresponding NEs are classified as a home-workplace pattern. Else are classified as other commuting patterns.

**Contribution.** In this research, focusing solely on network signals, accurate GPS-based anchor locations for individual movements like [13] can not be used. Thus in our case, grouping similar mobility motifs is much harder. We extract the most probably traversed route by user given locations of BTSs and shortest-fastest path analysis. Consequently the origin and the destination of user's trip are denoted by the first and last NEs respectively. As described in [12] we model NEs representing these motifs because individual's mobility can be characterized by their common mobility motifs. We group sets of NEs with similar trajectory patterns and select the most probable representing sets of NEs by a Bayesian

method. This selected set of NEs (model) helps impute the location of missing NEs and consequently improves trajectory reconstruction in weak signaling cases (see section V).

**3) Recognition of Mode of Transport.** Wang *et al.* [14] distinguished between 3 modes of transport: driving, walking, and using public transport, based on travel time. Wang *et al.* extracted travel time of a massive amount of passive NE traces for users traveling through the same origins and destinations in the radius of 500 meters. To discriminate mode of transport in cases with similar time interval, Doyle *et al.* [15] proposed a speed-based method, and distinguished modes of transport with significant speed difference. They compared the sequence of network cells for a path between a predefined origin and destination and categorized them as road users vs railway users.

**Contribution.** In this paper, we distinguish 6 modes of transport: driving by motorcycle/car, using a public transport such as bus, tram and train, riding a bicycle, and traveling on foot. We use an efficient mobility pattern recognition method which, beside achieving trajectory reconstruction, helps select the most probable mode of transport. We use a ranking procedure that nominates the transportation mode with highest rank as the most probable mode. The ranking procedure uses information such as graph models of geographical and public transport networks, bus/tram/train stops, and categorized average NE-based speed (see section VI).

### III. DATABASE

The database used in this work was collected using 3 test mobile devices carried by test users. The data contains 30 NE traces and the corresponding GPS traces of the trajectories. GPS traces is used to validate the performance of developed algorithms. The NEs are time-ordered collection of events, specific to each test mobile device and seen on 2G, 3G and 4G monitoring interfaces. The GPS traces are recorded by the same test mobile devices and apart from geo-locations they also contain annotations of the associated mode of transportation (*e.g.* walk, bicycle, car, bus, tram, and train). For each NE trace, we use following information: {*user-ID, time and date of NE trace collection, location of receptive BTSs (geographical latitudes and longitudes) and their time stamps, type of events, type of radio technologies*}. When collecting the GPS traces we record: {*user-ID, time and date of GP trace collection, mode of transport, GPS coordinates (geographical latitudes and longitudes) and their time stamps*}.

The statistics of traces used for this study is reported in Table I. Depending on the locations where these traces collected, we distinguish between three categories of areas: urban, suburban, and highway. The statistics include (a) the *average number* of GPS fixes, events and unique BTSs across traces. Unique number of BTSs is obtained after removing duplicated successive BTSs with similar latitudes and longitudes, (b) the *average duration* of trips, and the *average time interval between successive* GPS fixes, events, and unique BTSs in minutes, (c) the *distance interval between successive* GPS fixes, events, and unique BTSs in meters respectively.

TABLE I  
THE STATISTICS ON THE DATABASE OF 30 NE AND GPS TRACES.

Statistics	Urban	Suburban	Highway
No. GPS fixes	418	470	5569
No. of events	139	98	2386
No. of unique BTSs	8	8	77
Duration of trip (minutes)	11	7	54
Time int. of GPS fixes (minutes)	0.01	0.01	0.02
Time int. of events (minutes)	0.13	0.14	0.05
Time int. of unique BTSs (minutes)	1.96	1.36	0.83
Distance int. of GPS fixes (meters)	10	20	20
Distance int. of events (meters)	270	380	2540
Distance int. of unique BTSs (meters)	990	1080	2910

#### IV. MOBILITY PATTERN RECOGNITION AND TRAJECTORY RECONSTRUCTION

Humans or vehicles are bound to move in a geographical network. Thus, we analyze the embedded mobility patterns in this network and convert it into a graph. The edges and vertices of this graph represent the roads and road junctions, respectively. The graph edges are directed for one-way roads, or undirected otherwise. Weights are assigned to the edges based on heuristic criteria such as: the length of the road, the maximum speed limit or the importance of the road.

There exist many possible routes that connect an origin and a destination in a road graph. Finding all possible routes is a computationally demanding combinatorial problem. However, people usually travel by traversing the shortest or the fastest possible routes. Following this logic rather than extracting all possible routes, we focus on extracting and analyzing the set of shortest and fastest routes based on different modes of transport. These routes are extracted by applying Dijkstra algorithm [16] on corresponding (weighted) adjacency graphs. In this work, we extract these routes using <http://www.yournavigation.org/>, which is based on Open Street Map.

A “trajectory” is defined as a sequence of triples  $(X_i, Y_i, T_i)$ , where  $X_i, Y_i$  denote geographical latitude and longitude and  $T_i$  is a time stamp [17]. At each time stamp a GPS trace, a NE trace and a route are represented with GPS coordinates, location of BTSs (NEs), and route points respectively. To reconstruct the trajectory traversed by a mobile phone user based on NEs, we go through the following procedure.

For a given NE trace, we extract all plausible shortest and fastest routes between the locations of the first and the last network events. For each network event in the NE trace we find its corresponding point on candidate routes by minimizing pairwise Euclidean distances between that specific NE and all route points. We then calculate the posterior probability of each candidate route given NEs using Bayes theorem:

$$\begin{aligned}
 P(\text{Route}_j | \text{NEs}) &= \frac{P(\text{NEs} | \text{Route}_j) \times P(\text{Route}_j)}{P(\text{NEs})} \\
 &= \frac{\frac{1}{\|\text{NEs} - \text{Route points}_j\|_2} \times \frac{1}{\text{number of candidate routes}}}{\sum_{j=1} P(\text{NEs} | \text{Route}_j) \times P(\text{Route}_j)}, \quad (1)
 \end{aligned}$$

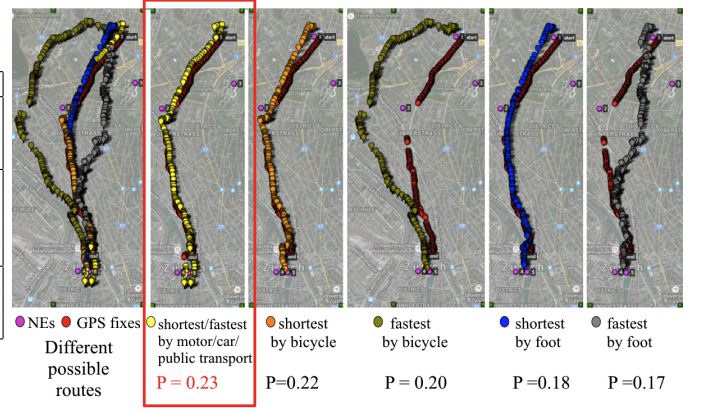


Fig. 1. Example of trajectory reconstruction algorithm for a NE trace.

where  $j = 1, \dots, \text{number of candidate routes}$ . The prior probability gives equal priority to each candidate routes. Likelihood  $P(\text{NEs} | \text{Route}_j)$  is computed using the inverse of Euclidean distance between given network events and representative route points.

A candidate route with the maximal posterior probability is the most probable representative route for the real trajectory traversed by the user.

Figure 1 depicts the result of trajectory reconstruction using aforementioned procedure. On the left, the real trajectory (GPS trace) in red and 4 different candidate routes between the first and last network events of a NE trace have been shown. In the other parts of figure you see each candidate route compared to the GPS trace and their posterior probability values  $P$ . The most representative route is selected as fastest/shortest by motor/car or public transport, which matches with real trajectory traversed by a tram. Note the minor difference between the most probable route and the GPS trace is due to the difference in the location of the first and the last NEs comparing to the first and last GPS fixes. The statistics of trajectory reconstruction for 30 different NE traces collected in urban area, suburban area and highways is reported in Table II. We compare the effect of using two different distances for computation of likelihood in equation 1: the inverse of Euclidean distance  $1/\|\text{NEs} - \text{Route points}\|$  and the inverse of Gaussian distance  $1/\exp\left(-\frac{\|\text{NEs} - \text{Route points}\|^2}{2\sigma^2}\right)$ , where  $\exp$  means exponential function and  $\sigma$  denotes the median distance between the NEs and the Route points. Beside trajectory pattern recognition rates, we also reported the mean distance  $\mu$  in meter and its standard deviation (std). Two distances are calculated: between the routes selected for given NE trace and the corresponding GPS trace, and between NEs and their corresponding GPS fixes. We obtained the best accuracy for highway traces, because the traces are recorded over longer time period, the spatial resolution of BTSs along highways is good, and users drive with high constant speed. The most difficult cases are NE traces that collected in the urban area because of a smaller number of NE traces collected in urban area, over short time intervals. Additionally, tall buildings

TABLE II  
TRAJECTORY PATTERN RECOGNITION RATES AND STATISTICS.

	Urban	Suburban	Highway	Total
<b>Recognition Rate %</b>				
Euclidean	75	100	100	<b>97</b>
Gaussian	75	73	100	87
<b><math>\mu(\text{Selected Route, GPS}) \pm \text{std in meters}</math></b>				
Euclidean	$378 \pm 75$	$103 \pm 87$	$127 \pm 19$	<b><math>203 \pm 152</math></b>
Gaussian	$540 \pm 114$	$288 \pm 173$	$127 \pm 19$	$318 \pm 208$
<b><math>\mu(\text{NE, GPS}) \pm \text{std in meters}</math></b>				
	$629 \pm 695$	1087	$1523 \pm 607$	<b><math>1082 \pm 452</math></b>

in city centers result in inaccuracies in GPS positioning, introducing additional error in the evaluation.

## V. MODELING MOBILITY PATTERNS

Users often repeat their mobility patterns, for instance commuting everyday from home to work place/university. They often travel through the same path with the same modes of transport. Suppose we collect the NE traces of a user while he/she was traversing the exact geographical route over a period of time. The question that raises here is whether his/her NE traces are similar as well, and if so, how we can model this repetitive trajectory pattern. The algorithm of modeling repeatedly traversed trajectory is following.

Suppose we have a set of NE traces representing the repeatedly traversed geographical route. Each  $\text{NE}_k$  trace is represented with a set of network events  $\text{NE}_{S_k}$ , where  $k = 1, \dots$ , number of NE traces. We group these NE traces if their associated GPS trace or selected route for trajectory reconstruction based on maximum posterior probability in equation 1 are similar.

We calculate the posterior probability of various NE traces for a given repeatedly traversed route using Bayes theorem as follows:

$$\begin{aligned}
 P(\text{NE}_{S_k} | \text{Route}) &= \frac{P(\text{Route} | \text{NE}_{S_k}) \times P(\text{NE}_{S_k})}{P(\text{Route})} \\
 &= \frac{\left( \frac{1}{\|\text{Route points} - \text{NE}_{S_k}\|_2} + \text{Coverage areas}_k \cap \text{Route points} \right)}{\sum_{k=1} P(\text{Route} | \text{NE}_{S_k}) \times P(\text{NE}_{S_k})} \\
 &\quad \times \frac{\text{Frequency NE}_{S_k}}{\text{Total frequencies}}, \quad (2)
 \end{aligned}$$

where the prior probability,  $\frac{\text{Frequency NE}_{S_k}}{\text{Total frequencies}}$ , means how many times the set of network events  $\text{NE}_{S_k}$  appeared in the considered set of NE traces. The coverage area of a BTS is calculated using its beam width angle and the minimum and maximum radius of its signal propagation.

A set of network events with maximum posterior probability models the repeatedly traversed trajectory pattern. In weak cellular signaling cases if user travel through the same path, this model enables one to find missing NEs and consequently reconstruct the trajectory given this selected set. Maximizing the inverse of Euclidean distance between a set of network events in a NE trace and route points allows selection of the network events nearest to the traversed route. Maximizing the

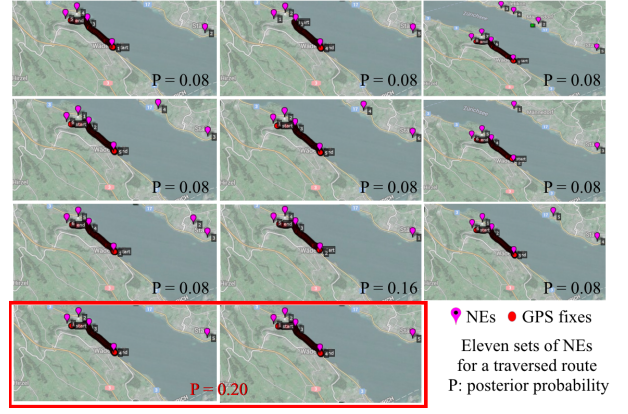


Fig. 2. Example of traversed trajectory modeling algorithm for different sets of NE traces.

intersection of BTS coverage areas with the traversed route filters out BTSs that are far from the traversed routes. The prior probability gives the importance to sets of NEs that repeated more often. Figure 2 depicts the result of modeling repeated trajectory patterns. For a geographical route traversed 11 times, shown with red markers, 11 corresponding NE traces are collected. The magenta markers represent NEs numbered in the order of their appearance in the trace. Maximizing the posterior probability of these eleven sets of NEs selects a set of NEs, which is as close as possible to the geographical route and has appeared more often. This set is highlighted with a red box in the figure 2.

## VI. MODE OF TRANSPORT RECOGNITION

Extracting the shortest and fastest routes based on different modes of transport affords a fast trajectory pattern recognition/reconstruction strategy, and at the same time allows recognizing the mode of transport. The selected candidate route for trajectory reconstruction shown in Figure 1 is a route that is traversed either by motor/car, or public transport. The exact mode of transport is recognized as follows: After the trajectory reconstruction step, we search in a heuristic 10 meter radius of the reconstructed route for bus, tram and train stops. For instance, on a road used by both tram and bus, finding more tram stops than bus stops results in assigned rank 1 to the mode of transport “tram”. We overlap the reconstructed route with the graphs of bus, tram, train and the road network. The graph with the highest correlation receives rank 1. We then compute the average NE-based speed of user traveling the reconstructed route using the latitude and longitude of the first and last network events and their time stamps:

$$\bar{V}_{\text{NE}} = \frac{Hav(\text{last NE} - \text{first NE})}{T_{\text{end}} - T_{\text{start}}} \quad (3)$$

where  $Hav$  denotes the Haversine distance (the distance between two points on a sphere, [https://en.wikipedia.org/wiki/Haversine\\_formula](https://en.wikipedia.org/wiki/Haversine_formula)) in radians between the last and first NEs.

A linear classifier assigned the average NE-based speed to the correct mode of transport category. We sum the ranks for



TABLE III  
MODE OF TRANSPORT RECOGNITION BASED ON AVERAGE SPEED (IN KILOMETERS PER HOUR).

Mode of Transport	$\bar{V}_{GPS} \pm \text{std}$	$\bar{V}_{NE} \pm \text{std}$
Car	$88.24 \pm 5$	$82.48 \pm 4$
Bus	$24.47 \pm 3$	$29.58 \pm 3$
Tram	$13.07$	$13 \pm 4$

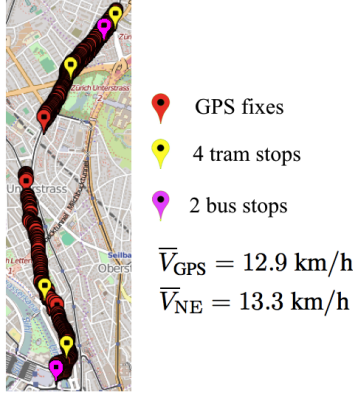


Fig. 3. Example of a GPS trace superimposed over an Open Street Map with relevant public transport stops along it. The relevant NE- and GPS-based speed approximations are shown.

each mode of transports (bus, tram, train, car/motor, bicycle, on foot). The mode of transport with highest assigned rank is recorded as the recognition result.

In equation 3 replacing the coordinates of the last and first GPS fixes results in the GPS speed. In Table III we compare the average NE-based speed,  $\bar{V}_{NE}$ , with the average GPS speed,  $\bar{V}_{GPS}$ , and their standard deviations over 30 traces used three different modes of transport. As you see values of  $\bar{V}_{NE}$  are very close to  $\bar{V}_{GPS}$  values, and are linearly separable.

Figure 3 shows an example of a GPS trace traversed by a tram. The GPS fixes, tram stops and bus stops are depicted with red, yellow, and pink markers respectively. This GPS trace is the example trace in Figure 1. The mode of transport for commuting through this route was narrowed down to motor/car, and public transport. We found more tram stops in the proximity of this route and the speed approximated using the NEs also suggests tram as the mode of transport.

## VII. CONCLUSION AND FUTURE WORK

In this paper we have analyzed the mobile phone users mobility based on network events, which are sparse both in terms of space and time. We have introduced a Bayesian approach that selects the most plausible route traversed by user from a set of shortest and fastest routes between the first and the last network events. This approach affords a fast trajectory pattern recognition/reconstruction strategy, and at the same time allows recognizing the mode of transport. We also have proposed a ranking procedure to further distinguish the modes of transport. To improve trajectory reconstruction in weak cellular signaling cases, we analyzed sets NEs corresponding

to repeatedly traversed geographical routes. We have modeled the traversed route with the most representative set of NES using a Bayesian approach. In future work we are planning to incorporate further networking evidence available via the cellular network, investigate clustering of network events, as well as validating the findings using a larger data set.

This work was supported by the grant from the Swiss Commission for Technology and Innovation (CTI).<sup>1</sup>

## REFERENCES

- [1] R. Becker, R. Cáceres, K. Hanson, S. Isaacman, J. M. Loh, M. Martonosi, J. Rowland, S. Urbanek, A. Varshavsky, and C. Volinsky, "Human mobility characterization from cellular network data," *Communications of the ACM*, vol. 56, no. 1, pp. 74–82, 2013.
- [2] D. Mountain and J. Raper, "Positioning techniques for location-based services (lbs): characteristics and limitations of proposed solutions," in *Aslib proceedings*, vol. 53, no. 10. MCB UP Ltd, 2001, pp. 404–412.
- [3] C. Ratti, D. Frenchman, R. M. Pulselli, and S. Williams, "Mobile landscapes: using location data from cell phones for urban analysis," *Environment and Planning B: Planning and Design*, vol. 33, no. 5, pp. 727–748, 2006.
- [4] R. Herring, A. Hofleitner, P. Abbeel, and A. Bayen, "Estimating arterial traffic conditions using sparse probe data," in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*. IEEE, 2010, pp. 929–936.
- [5] E. Saluveer and R. Ahas, "Using call detail records of mobile network operators for transportation studies," *Mobile Technologies for Activity-Travel Data Collection and Analysis*, p. 224, 2014.
- [6] F. Asgari, V. Gauthier, and M. Becker, "A survey on human mobility and its applications," *arXiv preprint arXiv:1307.0814*, 2013.
- [7] F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, S. Rinzivillo, and R. Trasarti, "Unveiling the complexity of human mobility by querying and mining massive trajectory data," *The VLDB Journal/The International Journal on Very Large Data Bases*, vol. 20, no. 5, pp. 695–719, 2011.
- [8] B. C. Csáji, A. Browet, V. A. Traag, J.-C. Delvenne, E. Huens, P. Van Dooren, Z. Smoreda, and V. D. Blondel, "Exploring the mobility of mobile phone users," *Physica A: statistical mechanics and its applications*, vol. 392, no. 6, pp. 1459–1473, 2013.
- [9] W. Andreas, W. Shangbo, J. Peter *et al.*, "Traffic congestion estimation service exploiting mobile assisted positioning schemes in gsm networks," *Procedia Earth and Planetary Science*, vol. 1, no. 1, pp. 1385–1392, 2009.
- [10] M. Fillekes, "Reconstructing trajectories from sparse call detail records," Master's thesis, Geographisches Institut der Universität Zürich, 2014.
- [11] G. Ranjan, H. Zang, Z.-L. Zhang, and J. Bolot, "Are call detail records biased for sampling human mobility?" *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 16, no. 3, pp. 33–44, 2012.
- [12] Z. Smoreda, A.-M. Olteanu-Raimond, T. Couronné *et al.*, "Spatiotemporal data from mobile phones for personal mobility assessment," *Transport survey methods: best practice for decision making*, vol. 41, pp. 745–767, 2013.
- [13] O. Järv, R. Ahas, E. Saluveer, B. Derudder, and F. Witlox, "Mobile phones in a traffic flow: a geographical perspective to evening rush hour traffic analysis using call detail records," *PloS one*, vol. 7, no. 11, p. e49171, 2012.
- [14] H. Wang, F. Calabrese, G. Di Lorenzo, and C. Ratti, "Transportation mode inference from anonymized and aggregated mobile phone call detail records," in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*. IEEE, 2010, pp. 318–323.
- [15] J. Doyle, P. Hung, D. Kelly, S. McLoone, and R. Farrell, "Utilising mobile phone billing records for travel mode discovery," 2011.
- [16] E. Dijkstra, "a note on two problems in connexion with graphs," *o numer*, 1959.
- [17] F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi, "Trajectory pattern mining," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2007, pp. 330–339.

<sup>1</sup>Swisscom strictly complies with all applicable registrations, in particular with the telecommunications law, and the data production initiative.